MLIB Documentation

The MLIB toolbox is available at:
https://de.mathworks.com/matlabcentral/fileexchange/37339-mlib-toolbox-for-analyzing-spike-data
or
https://github.com/maikstue/mlib-spike-data

The latest version of the accompanying paper is available at:
https://www.biorxiv.org/content/10.1101/2025.03.25.645246v1

For questions, bug reports, and suggestions for improvement, please contact:
Maik C. Stüttgen
Institute of Pathophysiology
University Medical Center Mainz
Duesbergweg 6
55128 Mainz
Germany
maik.stuettgen@uni-mainz.de

Table of Contents

Introduction 2
Function overview 3
Data file overview 4
Description of data files 6

Introduction

This document describes the use of the 21 functions contained in MLIB.

Additional information for each function is provided through each function's own documentation, as well as through the Function reference file (MLIB_Function_Reference.m). I suggest you open this m-file and run the individual sections therein, each of which demonstrates the use of one function.

In their respective sections, the functions are explained with reference to real or synthetic data. Real data was taken from several experiments in which we performed extracellular recordings from the brains of rats, mice, and pigeons. Most of this data is published:

- Kasties, N., Starosta, S., Güntürkün, O., & Stüttgen, M. C. (2016). Neurons in the pigeon caudolateral nidopallium differentiate Pavlovian conditioned stimuli but not their associated reward value in a sign-tracking paradigm. Scientific Reports, 6(October), 35469. https://doi.org/10.1038/srep35469
- Lengersdorf, D., Pusch, R., Güntürkün, O., & Stüttgen, M. C. (2014). Neurons in the pigeon nidopallium caudolaterale signal the selection and execution of perceptual decisions. European Journal of Neuroscience, 40(9), 3316–3327. https://doi.org/10.1111/ejn.12698
- Stoilova, V. v, Knauer, B., Berg, S., Rieber, E., Jäkel, F., & Stüttgen, M. C. (2020). Auditory cortex reflects goal-directed movement but is not necessary for behavioral adaptation in sound-cued reward tracking. Journal of Neurophysiology, 124(4), 1056–1071. https://doi.org/10.1152/jn.00736.2019
- van der Bourg, A., Yang, J.-W., Reyes-Puerta, V., Laurenczy, B., Wieckhorst, M., Stüttgen, M. C., Luhmann, H. J., & Helmchen, F. (2017). Layer-Specific Refinement of Sensory Coding in Developing Mouse Barrel Cortex. Cerebral Cortex, 27, 4835–4850. https://doi.org/10.1093/cercor/bhw280
- Vandevelde, J. R., Yang, J.-W., Albrecht, S., Lam, H., Kaufmann, P., Luhmann, H. J., & Stüttgen, M. C. (2023). Layer- and cell-type-specific differences in neural activity in mouse barrel cortex during a whisker detection task. Cerebral Cortex, 33, 1361–1382. https://doi.org/10.1093/cercor/bhac141
- Yeganeh, F., Knauer, B., Backhaus, R. G., Yang, J. W., Stroh, A., Luhmann, H. J., & Stüttgen, M. C. (2022). Effects of optogenetic inhibition of a small fraction of parvalbumin - positive interneurons on the representation of sensory stimuli in mouse barrel cortex. Scientific Reports, 12, 19419. https://doi.org/10.1038/s41598-022-24156-y

Function overview

Function name	Primary input data	Brief description
mcc[1]	spike waveforms from two or more clusters; spike time stamps	visualizes unit clusters and provides quality metrics to quantify spike sorting success
mcheck[1]	spike waveforms from one unit cluster; spike time stamps	visualizes one unit's waveforms and provides quality metrics
mcluster[1]	PSTHs from different units	clusters PSTHs from different units according to their similarity
mcsd	LFP data from a linear electrode array, averaged over multiple identical trials	performs a current source density analysis and plots the result
mcvar	spike time stamps from a single unit	computes the coefficient of variation for interspike intervals and the Fano factor for spike counts
mdimreduct[1]	PSTHs from different units	performs dimensionality reduction of spike responses from several units recorded under the same conditions
mep[1]	continuous local LFP data from a single or multiple electrodes	returns the evoked potential with error bars and confidence intervals
mlat[1]	spike and event time stamps	provides several measures of response latency
mmi[1]	spike counts obtained from the same unit under two different conditions	computes mutual information of spike counts
mmi_psych[1]	behavioral performance indices (hit rate and false alarm rate)	expresses behavioral performance in bits to compare with neuronal MI
mmultregress[1]	spike and event time stamps	performs a sliding-window multiple regression analysis of spike responses
mnspx	spike and event time stamps	provides a vector of spike counts in an interval around a series of events
mpsth	spike and event time stamps	constructs a PSTH and raster plot
mroc[1]	spike counts obtained from the same unit under two different conditions	computes the area under the ROC curve (AUROC) for two spike count distributions
msdf[1]	PSTH from one unit	constructs a spike-density function from a PSTH
mspectan	continuous LFP data	performs spectral analysis of LFP data
mtune[2,3]	spike and event time stamps	constructs and fits a tuning curve
mvecstrength	spike and event time stamps	computes vector strength of a spike response relative to a periodic signal
mwa[1]	spike and event time stamps	computes a 'window analysis', i.e., AUROC of a spike train in moving windows relative to a fixed reference event
mwa2[1]	spike and event time stamps	computes a 'window analysis', i.e., AUROC of two spike trains e.g., relative to two different stimuli
mwave	spike waveforms from a single unit	plots the average spike waveform and computes its characteristics (peak-to-trough duration etc.)

[1] requires the Statistics and Machine Learning Toolbox

[2] requires the Curve Fitting Toolbox

[3] requires the Signal Processing Toolbox

Data file overview

Most of the functions are illustrated using real data from experiments conducted by me and my colleagues (see references above). Table 1 lists the data files included with the MLIB toolbox; Table 2 provides a list of variables therein.

Table 1. Files used for illustration of selected MLIB functions, the variables they contain, the functions which are illustrated using the respective file, and the figure panels in the paper produced with them (if any).

No.	File name	Variables	Function(s)	Figure panels in paper
1	unitForMLIBTesting_1_rat	spike times (tspx) event codes (cevents) event times (tevents)	mcvar mnspx	none
2	unitForMLIBTesting_2_pigeon	spike times (tspx) event codes (cevents) event times (tevents) spike waveforms (adc)	mcheck mpsth msdf mwave	Figure 2a-h
3	unitForMLIBTesting_3_pigeon	spike times from two different channels (9 & 10, tspx_ch9 & _ch10) event codes (cevents) event times (tevents) waveforms from each channel (adc_ch9 & _ch10) cluster assignments for both channels' waveforms (clusters_ch9 & clusters_ch10) waveform sampling rate (fs)	mcc mcheck	Figure 2g-m
4	unitForMLIBTesting_4_rat	spike times (tspx) event codes (cevents) event times (tevents)	mlat mmi mwa	Figure 3a
5	unitForMLIBTesting_5_rat	spike times (tspx) event codes (cevents) event times (tevents)	mwa2	none
6	unitForMLIBTesting_6_rat	spike times (tspx) stimulus times (tstims) response times (tresponses) outcome times (toutcomes) stimulus type (stims) response type (responses) outcome type (outcomes)	mmultregress	Figure 3c
7	data4mcluster	concatenated peri-stimulus time histograms for 98 single neurons both in units of firing rates (app_psths) and z-scored firing rates (zpsths)	mcluster	none
8	data4mcsd	average LFP responses from 20 electrode channels over 80 ms (mean_fp)	mcsd	Figure 5
9	data4mdimreduct	spike rates of 300 units relative to presentation of either of two stimuli (nspx_s1, nspx_s2) in 17 time windows (winpos, from -0.4 to +0.4 relative to stimulus onset, shifted in steps of 50 ms); each spike rate was computed over 100 ms (winsize)	mdimreduct	Figure 4
10	data4mep	continuous field potential trace from a single electrode (fp) sampled at 1 kHz (fs); 26 stimuli were presented in pseudorandom order (cevents) every 4 s (event timestamps in tevents)	mep	none
11	data4mspectan	continuous EEG recording from a single electrode (eeg) sampled at 1 kHz (fs)	mspectan	none

Table 2. Variables in the above data files and their contents.

Variable name	Type	Content
adc	array	waveform samples rows are waveforms columns are samples
app_psths	array	peri-stimulus time histograms for 98 units (rows) for five 5-s stimuli; firing rates were calculated over non-overlapping 0.2-s-analysis windows, yielding 25 bins per PSTH; PSTHs for the five stimuli are appended for each unit
cevents	vector	event codes (integers), specific for each data file; also see tevents for times
eeg	vector	continuous EEG recording
fp	vector	continuous LFP recording
fs	scalar	waveform sampling rate in Hertz (applies to spike waveform and LFP data)
mean_fp	array	average LFP recorded from 20 electrodes (rows) over 80 ms (columns)
stims	vector	stimulus sequence for unitForMLIBTesting_6_rat.mat; stimulus timestamps are contained in variable tstims
tevents	vector	event timestamps, all in seconds; there are different types of events, the key is in variable cevents
tspx	vector	spike timestamps, all in seconds
tstims	vector	timestamps of stimuli in vector stims in unitForMLIBTesting_6_rat.mat
winpos	vector	positions of analysis window (in s) used for spike rate calculations in data4dimreduct
winsize	scalar	size of analysis window (in ms) used for spike rate calculation in data4dimreduct
zpsths	array	as app_psths, but normalized (z-scored)

Description of data files

<u>Spike recordings from the pigeon nidopallium caudolaterale</u>

Pigeons were implanted with multi-electrode microdrive arrays in the nidopallium caudolaterale (NCL), an associative area of the avian brain which is believed to serve functions similar to mammalian prefrontal cortex. The animals were performing a visual discrimination task for food rewards (files #2 and #3, Lengersdorf et al., 2014). Briefly, the birds initiated a trial by pecking at a rectangular response key (the 'center key'). This was followed by the presentation of one of six different visual stimuli (shades of gray, ranging from dark to bright) for 1 s. Following presentation of either of the three darker stimuli, the animal should peck at a response key on the right of the center key, and for the three brighter stimuli, it should peck at a response key on the left. If the response was correct, 2 s access to food was granted according to a probabilistic schedule. Incorrect responses were punished with 2-s timeouts.

File #7 (data4mcluster.mat) is from a sign-tracking experiment (Kasties et al., 2016). On each trial, pigeons were presented with one of five visual stimuli (colored textures) which were assigned with different reward magnitudes (starting from zero reward). Each stimulus was presented for five seconds and 25 times per session.

For all files, electrode signals were amplified 1000x, bandpass-filtered from 0.5 to 5 kHz and subsequently digitized at a resolution of 16.67 kHz. Full methodological details can be found in the above-mentioned papers.

File #2: unitForMLIBTesting_2_pigeon.mat

This file contains 853 waveforms from a single NCL unit (adc), each with 32 samples. Spike time stamps are contained in tspx (in seconds). The variable tevents includes time stamps of experimental events, the variable cevents specifies which event time stamp belongs to which experimental event.

The event codes are as follows:

1: trial initiation, 2-7: six visual stimuli (shades of gray, from dark to bright), 8: begin of choice phase, 9: food reward, 10: food omission, 11: punishment, 12-14: key pecks to the left, center, and right pecking keys, respectively.

File #3: unitForMLIBTesting_3_pigeon.mat

Similar to unitForMLIBTesting_2_pigeon.mat, this file contains waveforms from two electrodes (channels 9 & 10) in the same session (variables adc_ch9 and adc_ch10, respectively). Both channels contain ALL waveforms detected on a given channel. Waveforms were sorted into multiple clusters using the Spike2 software. Which waveforms was assigned to which cluster is coded in the variables clusters_ch9 and cluster_ch10. For both channels, cluster 0 contains noise, artifacts, and waveforms which could not be unambiguously assigned to any of the other clusters. clusters_ch9 contains two spike clusters (coded 1 and 2), clusters_ch10 contains three spike clusters (coded 1, 2, 3). The time stamps of all waveforms are contained in tspx_ch9 and tspx_ch10.

Events and their time stamps are coded as described above, but here codes represent:

97: trial initiation, 98-103: six visual stimuli (shades of gray), 104: begin of choice phase, 105: food, 106: food omission, 107: time-out punishment, 108-110: pecks to the left, center, and right pecking keys, respectively.

File #7: data4mcluster.mat

This file contains peri-stimulus time histograms (PSTHs) from 98 single units recorded from the pigeon NCL while the animal was sign-tracking. For each of five visual stimuli (presented for 5 s), a PSTH relative to stimulus onset was constructed. PSTH resolution is 200 ms, i.e., each PSTH consisted of 25 data points. The variable app_psths contains five concatenated PSTHs (columns, sorted in ascending order of associated reward magnitude) for 90 single neurons (rows). The variable zpsths contains the same data but after z-scoring all PSTHs to control for baseline firing rate.

<u>Spike recordings from the rat cortex</u>

Rats were implanted with multi-electrode microdrive arrays into the left or right auditory cortex (AC) or medial prefrontal cortex (MFC). The animals were performing a two-stimulus conditional discrimination task for water reward from (Stoilova et al., 2020). Briefly, the rats initiated a trial by nose-poking into a central response ports for 400 ms. Subsequently, one out of two 70-ms noise bursts differing in frequency band was presented. Animals had to nose poke into another port located on the right after a low-pitch noise burst and into a nose port on the left after a high-pitch noise burst. Correct responses were rewarded with 60 µl of water delivery at the respective left or right choice port on a probabilistic basis (50% of correct trials). Incorrect responses were punished with 4-s timeouts.

For all files, electrode signals were amplified 1000x, bandpass-filtered from 0.5 to 5 kHz and subsequently digitized at a resolution of 20 kHz. Full methodological details can be found in the above-mentioned paper.

Files #1, #4, #5: unitForMLIBTesting_1_rat.mat, unitForMLIBTesting_4_rat.mat, unitForMLIBTesting_5_rat.mat

These files contain three variables: spike times (tspx), event times (tevents), and event codes (cevents). All timestamps are in seconds. Event codes signify which event occurred at which time: for purposes of illustration, it suffices to say to event 1 corresponds to the onset of stimulus 1 and event 2 corresponds to the onset of stimulus 2.

This data is from Stoilova et al. (2020).

File #6: unitForMLIBTesting_6

Same as above, but recordings are from medial prefrontal cortex, and time stamps of several trial events are provided. The variable tstims contains time stamps of stimulus presentations (stimulus 1 or stimulus 2, contained in stims), time stamps of responses (tresponses, nose poking into the right or left port contained in variable responses, 1 or 2, respectively), and time stamps of trial outcomes (toutcomes, events in outcomes are 1: correct, reward presentation, 2: incorrect, no consequence).

This data is unpublished.

File #9: data4mdimreduct

This file serves to illustrate the use of mdrimreduct.m to plot neural population trajectories (e.g., (Saxena & Cunningham, 2019). The file contains data from 300 units recorded in the medial frontal cortex of rats performing the two-stimulus conditional discrimination task described above. The arrays nspx_s1 and nspx_s2 contain spike counts from the 300 units (rows), obtained by shifting a 100-ms analysis window (winsize) in steps of 50 ms, starting 400 ms before stimulus onset and the following 400 ms (17 data points per unit, winpos). nspx_s1 gives spike counts relative to trials in which stimulus 1 was presented, accordingly for nspx_s2. Only correct trials were included.

This data is unpublished.

<u>Local field potential recordings from mouse barrel cortex</u>

File #8: data4mcsd.mat

This file serves to illustrate the use of mcsd.m for performing a current-source density (CSD) analysis. A four-shank silicon probe featuring 4*20 linearly spaced electrodes was inserted perpendicular to the cortical surface of an anesthetized juvenile mouse (P11). Inter-electrode distance was 50 µm. Individual whiskers deflected every few seconds during recording.

The file contains a single variable, mean_fp. The 20 rows represent 20 electrodes from a single shank located in the principal barrel column (i.e., the column corresponding to the whisker that was deflected). Columns represent the local field potential averaged over multiple stimulus presentations, sampled at 1000 Hz, for 80 ms.

This data is taken from van der Bourg et al. (2017) and used with friendly permission of authors Jenq-Wei Yang and Heiko J. Luhmann.

File #10: data4mep.mat

A two-shank silicon probe featuring 2*32 linearly spaced electrodes was inserted perpendicular to the cortical surface of an anesthetized adult mouse. Inter-electrode distance 25 µm, sampling rate fs was 1000 Hz. Overall, 26 different types of whisker deflections were applied (differing in amplitude, stimulated whisker, with or without laser illumination to inhibit fast-spiking interneurons). Every stimulus was applied 24 times, in pseudorandom order. Stimulus codes are in variable cevents, their timestamps in variable tevents. For present purposes, I have extracted the continuous recording from a single electrode located in layer 2/3 (variable fp).

This data is taken from Yeganeh et al. (2022).

<u>Electroencephalography recordings from the human brain</u>

File #11: data4mspectan.mat

EEG was recorded continuously for 78.67 seconds over the occipital cortex of a male human subject using the Backyard Brains Heart & Brain Spiker Box with the Spike Recorder Software (variable eeg). Sampling rate fs was 1 kHz, voltage is expressed in arbitrary units. The subject was asked to open and close his eyes about every 10 seconds.

This data is unpublished.